Combining Bayesian Optimization and Neural Network to Optimize the Plasma Wakefield Acceleration

Jiabao Guan,[1,2] Chang You,[3] Yuancun Nie,[1,2,*] Guoxing Xia,[4,5,†] and Jike Wang[1,2,‡]

[1]*The Institute for Advanced Studies, Wuhan University, Wuhan, 430072, China*
[2]*Advanced Light Source Research Center, Wuhan University, Wuhan 430072, China*
[3]*School of Physics and Technology, Wuhan University, Wuhan 430072, China*
[4]*Department of Physics and Astronomy, University of Manchester, Manchester, United Kingdom*
[5]*Cockcroft Institute, Daresbury, Warrington, WA4 4AD, United Kingdom*

The computational cost of finding the optimal design of plasma wakefield acceleration (PWFA) is usually very demanding due to many variables involved. Herein, we have developed a novel framework which combines Bayesian Optimization (BO) with neural network (NN), to replace computationally expensive simulation software and provide a more efficient way for the optimization process. In order to verify this framework, the AWAKE Run 2 experiment at CERN is used as an example. In the framework we constructed, the coefficients of determination (R^2) of NN reaches above 0.99, and the time-to-solution reduces to a factor of 35.6. For the first time, BO combined with NN is successfully applied to optimize PWFA and significant improvements have been demonstrated. The framework established here in principle can also be extended to the optimization of other particle accelerations.

Keywords: Plasma wakefield acceleration, Bayesian optimization, Neural network

I. INTRODUCTION

For traditional radio-frequency (RF) based accelerators, it is preferable to increase the feed-in power of the RF cavity to increase the acceleration gradient. However, a continuous increase of the cavity voltage will lead to the electrical breakdown of cavity materials, which poses a fundamental limit for conventional accelerators. On the other hand, the plasma-based accelerators, either driven by a ultrashort and intense laser pulse or a relativistic particle bunch, can sustain GV/cm acceleration gradient[1–3], i.e. more than three orders of magnitude compared with the conventional accelerators. Therefore, the plasma-based accelerators can greatly reduce the footprint of future accelerators and their associated construction and operation cost. It is envisaged that the plasma-based accelerators will find many applications in providing high quality particle beams for free electron lasers [4–7], fixed target experiments and energy frontier colliders [8–11].

In the process of design and optimization of a plasma accelerator, many variables, including plasma parameters, driving beam parameters and witness parameters should be taken into account simultaneously. To obtain a high quality beam, it is often necessary to continuously optimize these variables, such as brute force search or use of optimization algorithms. Brute force search requires a lot of computing resources and time, which increases exponentially with the number of variables [12]. Once the number of variables is too large, the calculation time required will be unacceptable. The use of optimization algorithms, such as genetic algorithm [13, 14], particle swarm optimization algorithm [15] and Bayesian Optimization (BO) algorithm [16, 17], can reduce the time to find the optimal solution to a certain extent [18]. But if there are multiple optimization problems in the same variable range, the algorithm optimization process needs to be repeated for each optimization problem, which also brings an increase in time and computing resources. The neural network(NN), however, can solve this problem very well. The training set data is generated through the simulation software. The model generated by the NN training can be used to replace the simulation software for prediction when the prediction accuracy is sufficient. When there are multiple variables, it is no longer necessary to use simulation software to calculate, and the results can be predicted through the trained neural NN instead.

BO is an efficient machine learning tool for solving multivariable, complex objective functions [16, 17]. It has been successfully applied to several laser-driven electron acceleration simulations [19, 20] and experiments [21, 22]. The calculation of the objective function value in the BO process is often obtained through simulation data from software or data collected from experiments, which requires a lot of computing resources and time. In this paper, in the calculation of the objective function value in the BO process, the predicted value from the NN is used instead of the calculated value of the simulation software. Since the prediction time using the NN is much less than the calculation time using simulation software, this greatly reduces the calculation time and computing resource requirement. In addition, when different physical problems need to be optimized simultaneously, the use of NN predictions can greatly reduce the computing resources [23]. In this paper, we introduce this novel framework that combines BO and NN, by taking AWAKE Run 2 experiment at CERN as an example, to demonstrate the effectiveness of this method. Section II introduce the physical problems of AWAKE Run 2 and methods used. Work flow and optimization results are given in Section III and IV. The discussion and conclusion are presented in Sections V and VI, respectively.

* Corresponding author, nieyuancun@whu.edu.cn
† Corresponding author, guoxing.xia@manchester.ac.uk
‡ Corresponding author, jike.wang@whu.edu.cn

II. METHODS

Physical problems and parameter settings in AWAKE

The Advanced WAKEfield (AWAKE) experiment at CERN is the world first and only proton-driven plasma wakefield acceleration (PWFA) experiment [24–27]. The first phase of AWAKE experiment (2016-2018) has successfully demonstrated the self-modulation of a long proton bunch and acceleration of an externally injected low energy (∼18.5 MeV) electron bunch up to 2 GeV in a single stage of 10 m long plasma channel [37–39]. The current experiment, so-called AWAKE Run 2 [28], aims to accelerate an electron bunch up to ∼10 GeV with a narrow energy spread and sufficiently low emittance for high-energy physics applications [8, 9].

The AWAKE Run 2 experiment uses the CERN SPS 400 GeV proton beam with an RMS length of 6-12 cm, after self-modulation in a plasma [29–31], multiple micro-bunches are formed as the driving beam for the PWFA [32]. In the previous study, the toy model [33] is used to adjust the charge of the proton beam to make the plasma wakefield reach the desired acceleration gradient. The proton beam energy, bunch length and beam radius are 400 GeV, 40 μm, and 200 μm, respectively. In addition, the mismatch of the beam radius can lead to a significant beam emittance growth [33, 40]. The matched radial beam size for a Gaussian beam in the plasma ion column is defined by [33, 41]:

$$\sigma_r = \left(\frac{2\varepsilon_0}{\gamma_0 k_p^2}\right)^{\frac{1}{4}}, \tag{1}$$

where k_p=$\frac{\omega_p}{c}$ is the plasma wave number, ε_0 is the initial geometric emittance, γ_0 is the initial mean Lorentz factor, ω_p and c are the plasma oscillation frequency and the speed of light, respectively. ω_p is determined by the plasma density n_0 and is given by

$$\omega_p = \sqrt{\frac{n_0 e^2}{m_e \epsilon_0}}, \tag{2}$$

where e, m_e and ϵ_0 are the electron charge, electron mass and the permittivity of free space, respectively. According to Eqs. (1) and (2), the calculated plasma density is 7×10^{14} cm^{-3}.

Fig. 1 gives the schematic diagram of AWAKE Run 2 setup. It uses a 400 GeV proton beam to pass through a 10 m long plasma to excite the wakefield. The electron beam is injected externally, and the appropriate acceleration phase is set by adjusting the position between the driving beam and the witness bunch. The electrons will be captured by the plasma wakefield and accelerated. In the simulation of AWAKE Run 2 experiment, the toy model is employed. The parameters of the driving bunch and the plasma are fixed and the initial energy of the witness bunch is set to 150 MeV [28, 34]. The optimal solution is obtained by optimizing the acceleration distance and five parameters of the witness bunch, i.e., the bunch charge, bunch length, emittance, energy spread and its position from the driving bunch. The variation ranges of the six variables are shown in Table 1, and the witness bunch position is the distance between the witness bunch and the driving bunch, in $k_p{}^{-1}$. $2\pi/k_p$ is the wavelength of the plasma wakefield.

TABLE 1. Variables in the optimization problem and the corresponding range of variation

Variables	Variation range
Acceleration distance	0 - 10 m
Witness bunch charge	50 - 300 pC
Witness bunch length	40 - 120 μm
Witness bunch emittance	4 - 10 mm·mrad
Witness bunch position	5.8 - 6.7 $k_p{}^{-1}$
Witness bunch energy spread	0.1% - 0.5%

Neural Network

In a fully connected NN, each neuron is connected to all neurons in the previous layer, so a weight matrix is needed to describe these connections. This weight matrix is called the kernel, which determines how the output of neurons in the previous layer is combined into the input of the neurons of the current layer. In addition, each neuron also has a learnable bias parameter, which is a constant added to the input of all neurons in the current layer to adjust the output of the model. The output of the neuron is activated only when the input to the neuron exceeds a certain threshold, which is controlled by a bias parameter. As shown in Fig. 2, each network layer contains kernel and bias parameters. During the training process, the network will automatically learn these parameters so that the prediction results of the model are as close as possible to the true value. The goal of optimization is to obtain the best beam quality, including witness bunch energy, energy spread, emittance and charge after being accelerated in plasma wakefield in a 10-meter-long plasma. Therefore, when training the NN for this physics problem, the NN should contain one input layer with six neurons and one output layer with four neurons. After optimization, the number of hidden layers of the NN is set to four, and each hidden layer contains 100 neurons. Within the range of variables, 2000 variable combinations were randomly sampled, and the corresponding four parameters of the witness bunch after 10-meter acceleration are calculated by the simulation software LCODE [35, 36]. The 2000 events obtained from LCODE simulation are used to train the NN, of which 1600 events are used as the training set, 200 events are used as the verification set, and the remaining 200 events are used as the test set.

BO algorithm

BO is an efficient algorithm for finding the optimal solution. It will predict the probability of the target value corresponding to the variable space based on the data of all previous iterations, then select the variable corresponding to the

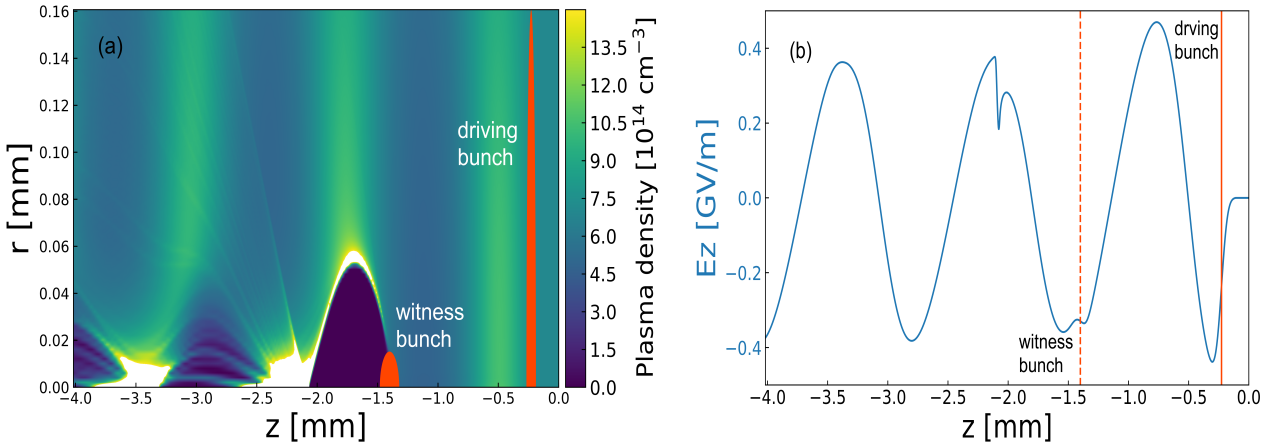

Fig. 1. Schematic diagram of proton beam driven plasma wakefield acceleration. (**a**) The plasma charge density distribution during the PWFA process. The black part is the bubble region generated by the proton beam driving the plasma wakefield. The left and right orange parts are the relative positions of the witness bunch and the driving bunch in the plasma, respectively. (**b**) The blue curve is the electric field distribution of the plasma wakefield. The orange dashed and solid lines show the relative positions of the witness bunch and the driving bunch in the plasma wakefield, respectively.

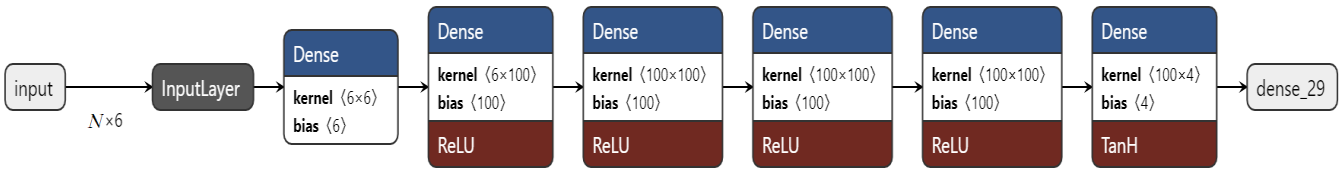

Fig. 2. Schematic diagram of a fully connected NN. The figure shows the dimensions of the internal kernel and bias matrix of the NN. The hidden layer uses the Rectified Linear Unit (ReLU) activation function, and the output layer uses the Hyperbolic Tangent (TanH) activation function, which maps the input of the neuron to the output. "$N\times6$" indicates the dimension of the input tensor. "N" is the number of samples in the data set, indicating the uncertainty of a certain dimension, which allows the NN to accept any number of samples as input and adapt to different input data set sizes. "6" is the number of eigenvalues for each sample, which is the 6 variables mentioned above.

predicted optimal target value as the next generation, and continuously get a better solution. In the AWAKE Run 2 optimization problem, J. P. Farmer and L. Liang et al. develop a single figure of merit that can give rise to constraints on both the tunability and stability of the initial witness bunch parameters [40]. After the NN is trained, the BO combined with NN framework can be used to optimize according to different optimization objectives. Here we take a four-dimensional target as an example, including the energy (E), energy spread (δ), emittance (ε) and charge (Q) of the electron bunch. Therefore, when using BO, the f value of Eq. (3) is used as the single-objective optimization. The larger the f, the better the beam quality. In this BO combined with NN framework, different optimization objectives can be selected according to different research problems.

$$f = \frac{E(GeV)\cdot Q(pC)}{\delta(\%)\cdot\varepsilon(mm\cdot mrad)}. \tag{3}$$

III. WORK FLOW

As shown in Fig. 3, we first use the plasma wakefield simulation software LCODE to randomly generate 2000 events as data set, and then randomly divide the data set into training set, verification set and test set according to the ratio of 8:1:1. We then build a fully connected NN with 6 inputs and 4 outputs. Through inputting the training set and verification set data into the NN, we can continuously adjust the NN parameters, including the number of hidden layers, learning rate, batch size and the number of neurons in the hidden layer. Finally, the parameters combination with the highest prediction accuracy of the test set data are selected as the trained NN.

After the NN is trained, the BO algorithm is used to optimize the problem. During the optimization process, when the initial sampling point is generated and the goal value of the new observation point is calculated, the objectives value predicted by the NN is used to calculate the goal value. For different optimization objectives, the convergence iterations of the optimization process is also different. In order to ensure the convergence of BO, a higher iteration threshold for the end of BO is set in the workflow. When the number of

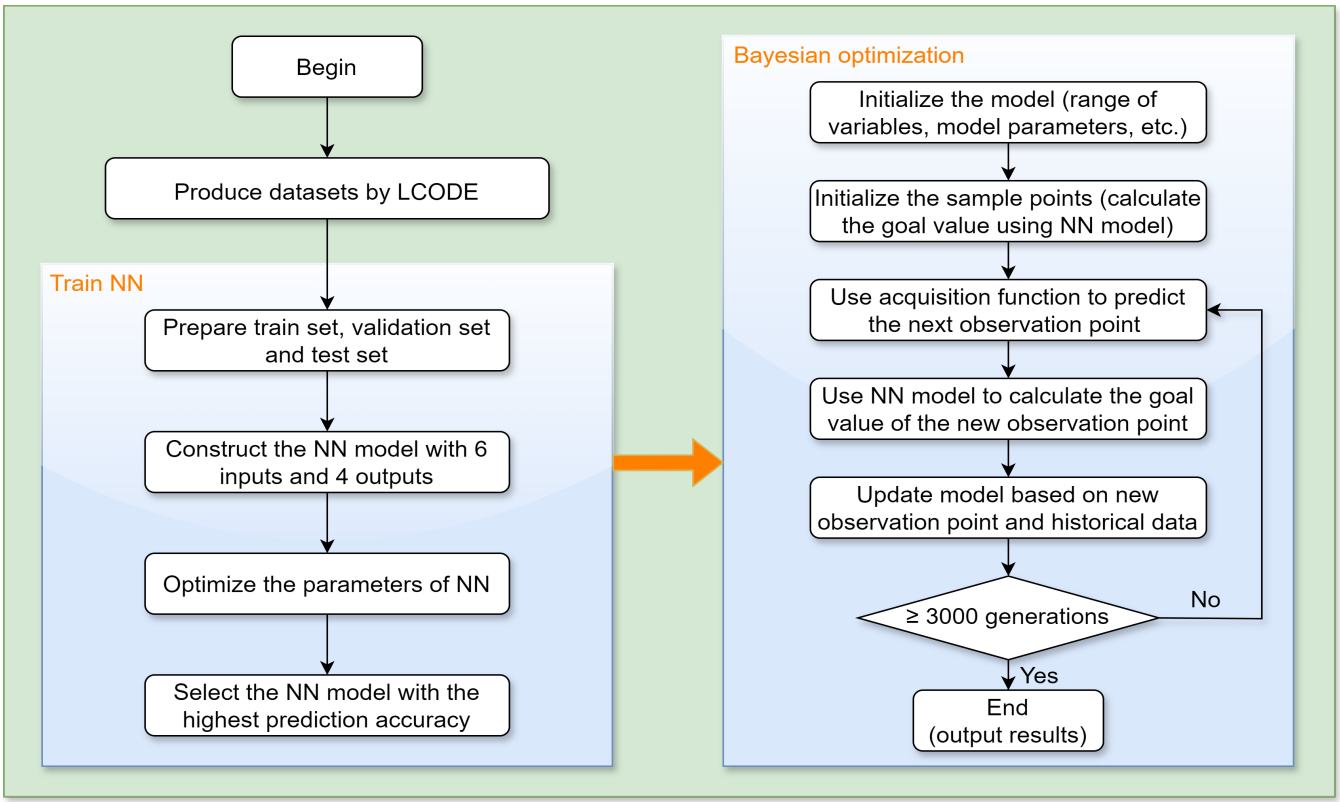

Fig. 3. Optimizing proton-driven PWFA using BO with NN. The blue area on the left is the training NN workflow, and the blue area on the right is the BO process.

iterations reaches 3000, the optimization iteration ends, and the optimal solution parameters and goal values are output.

IV. RESULTS

Training Neural Network

By optimizing the parameters of the NN including the number of hidden layers, learning rate, number of neurons and epoch, a feed-forward NN for the AWAKE Run 2 problem is established. Suppose a data set includes a total of y_1 to y_n observations, and the corresponding model predicted values are f_1 to fn respectively. The coefficients of determination (R^2) is calculated by the following:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - f_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \ \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \quad (4)$$

Through the evaluation of the R^2 parameter, the accuracy of NN prediction, that is, the R^2 between the target value predicted by the NN in the calculation test set and the target value calculated by LCODE, is close to 1, indicating that the prediction accuracy is high. The R^2 values of the six objectives are shown in Table 2, and the average R^2 value reaches 0.9975.

TABLE 2. R^2 values for the four objectives in the test set.

Objectives	Energy	Spread	Emittance	Charge	Average
R^2 value	0.9999	0.9989	0.9917	0.9996	0.9975

There are 200 events in the test set, and the comparison between the value predicted by the NN and the value calculated by LCODE is shown in Fig. 4. The beam energy, energy spread, emittance and bunch charge are compared from top to bottom. The blue point is calculated by LCODE and the orange point is predicted by the NN. The pink points are the error values for each individual between calculation and prediction. The closer the prediction is, the better the NN prediction works.

BO with NN

We employ the BO to optimize the AWAKE Run 2 problem, and use the f value as the single-objective optimization. The larger the f value, the better the beam quality. As shown in Fig. 5, after 3000 iterations, the highest f-value is at generation 2860.

The f-value of the 2860^{th} generation is the highest, which is the optimal solution obtained by the optimization. At this time, the corresponding variables, namely the bunch

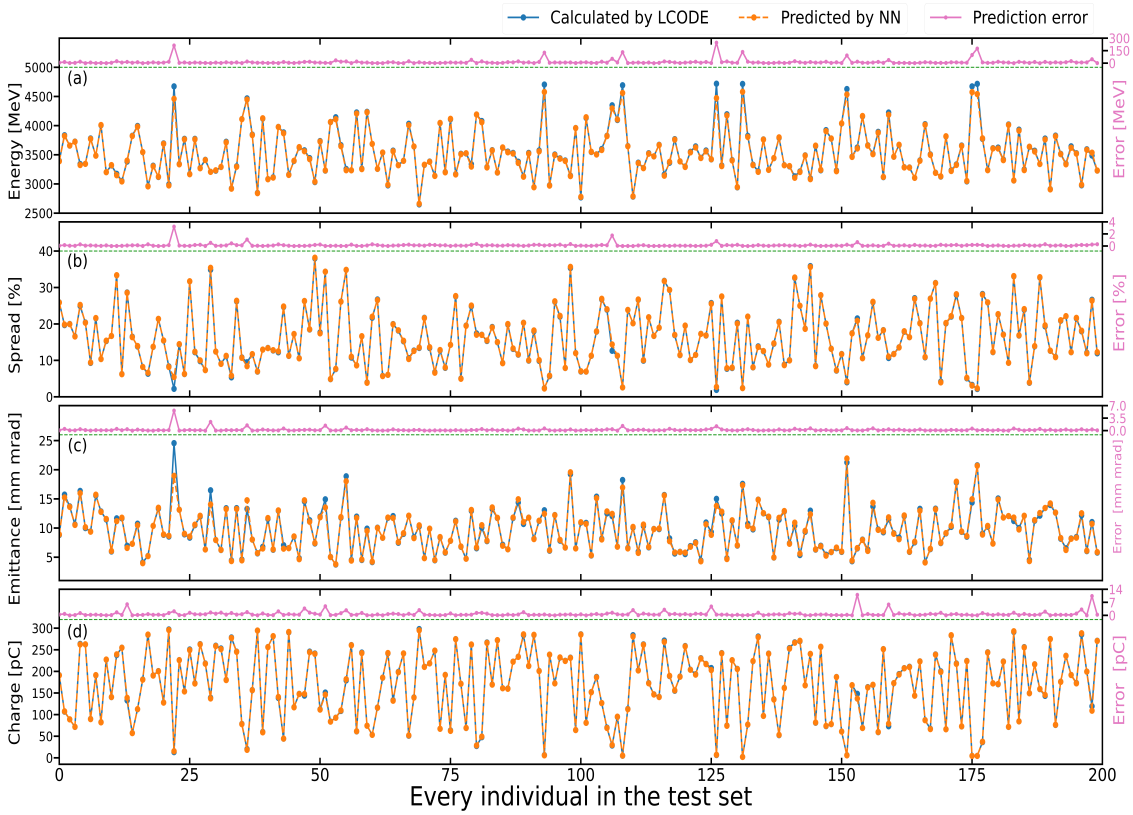

Fig. 4. The comparison between the NN prediction value and the LCODE calculation value of each data individual in the test set containing 200 events. (a), (b), (c) and (d) are the output beam energy, energy spread, emittance and charge respectively. The blue points are the calculated value of LCODE, the orange points are the predicted value of the NN model. The pink points are the error values for each individual between the value predicted by the NN and the value calculated using LCODE.

charge, bunch length, emittance, position and energy spread are 135.13 pC, 41.49 μm, 2.00 mm·mrad, 5.95 and 0.17% after a 10 m accelerate distance. The corresponding target values are as shown in Table 3. The first row is the result predicted by the NN, and the second row is calculated by LCODE to verify the accuracy of the prediction. It can be seen that each target value predicted by the NN is close to the result obtained by the LCODE calculation and verification, indicating that the prediction accuracy of the NN is high.

In the BO process, there is a tendency to approach the optimal solution, so the convergence trend of each variable can be judged by the distribution density of the polylines in the parallel axis coordinate graph. As shown in Fig. 6, the denser the polyline is, the more likely it produces a better solution within this dense range.Through the calculation of hyperparameter importance in the BO process, as shown in Fig. 7, we find that the energy spread has the highest percentage in the importance, indicating that the initial energy spread of witness bunch plays a more important role for optimizing the goal. This means, in the process of optimization to find the optimal solution, initial energy spread is the main variable, and changing the initial energy spread has the greatest impact on the goal value, which can provide empirical help for our optimization problems.

V. DISCUSSION

Here we compare the time-to-solution and computational resources of the two strategies. Strategy 1 is to use BO, and LCODE is used for calculation during the optimization process. Strategy 2 is the approach discussed above, which uses framework (BO combined with NN) for optimization. The advantage of using BO combined with NN compared to the approach of BO and simulation software LCODE in terms of time is evaluated from Bellotti's method [23]. The time required for different processes such as using LCODE to calculate once, as well as the number of computer cores used by different processes and their corresponding values are shown in Table 4. Time-to-solution is the time required to find the optimal solution. Because the time spent using different numbers of cores is different, therefore, when calculat-

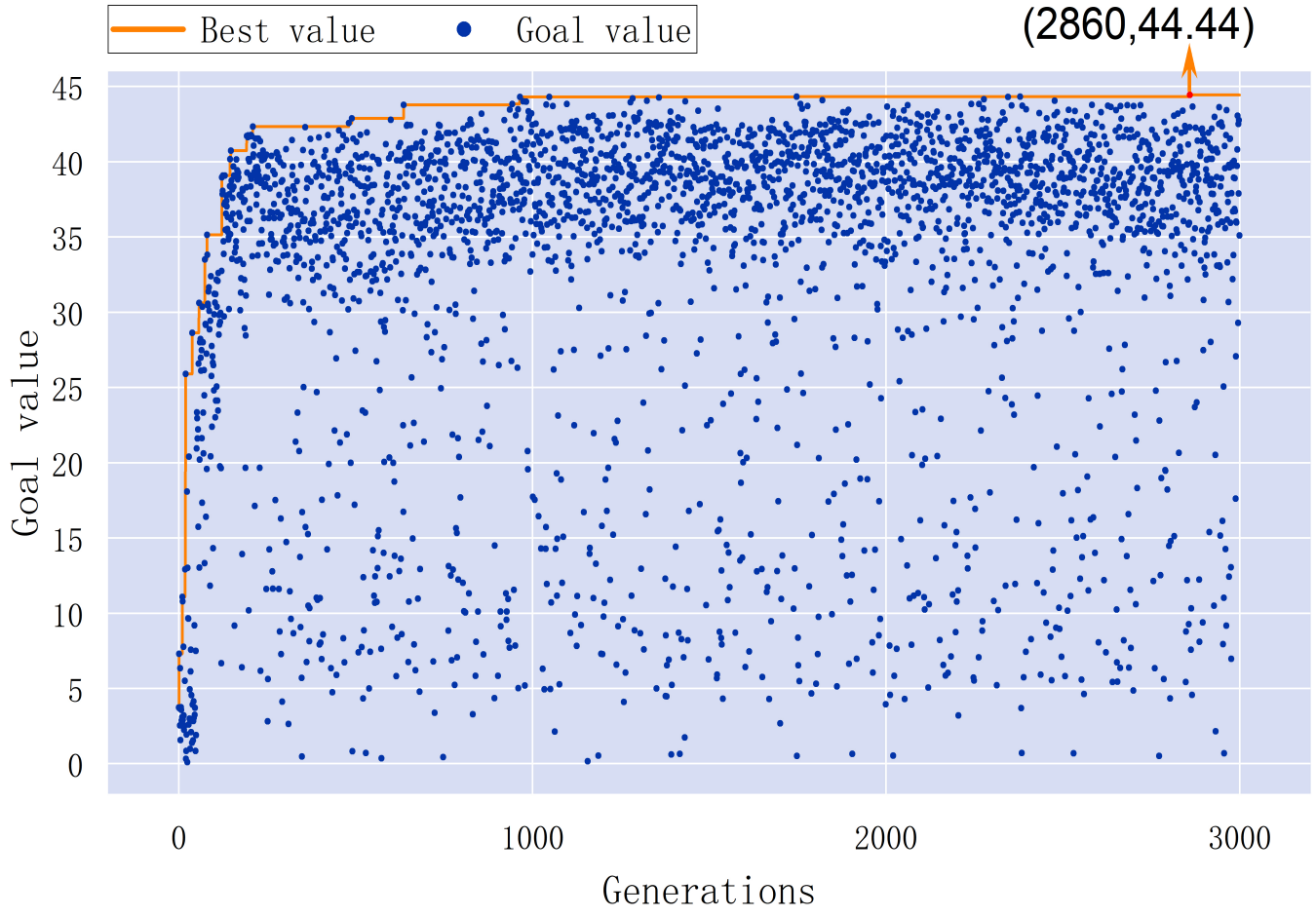

Fig. 5. BO iteration figure. The blue points represent the f-values calculated for each iteration, and the points on the orange line represent that the f-values of each new iteration are higher than the f-values of all previous generations.The red point indicates the highest goal value, that is, the optimal solution, which is the 2860^{th} generation in the optimization process.

TABLE 3. The NN prediction value of the optimal solution and the verification value calculated by LCODE.

	Energy(MeV)	Spread(%)	Emittance(mm·mrad)	Charge(pC)	f-value
NN	3509.85	2.94	3.63	135.34	44.44
LCODE	3529.50	2.91	3.52	135.10	46.55

ing time-to-solution, it is assumed that computing resources are unlimited, i.e., there are infinite cores that can be used. The time t_{LCODE} required to find the optimal solution using the strategy of BO with simulation software is shown in Eq. (5). Eq. (6) is the time t_{FW} to find the optimal solution using the framework. The ratio of the two shown in Eq. (7) gives the time saved by using the neural NN instead of the simulation software. In addition, the usage of computing resources of the two strategies is shown in Eqs.(8) and (9) respectively, considering the number of cores used for LCODE calculation and training NN. c_{LCODE} and c_{FW} are the computing resources used by strategies 1 and 2, respectively. Eq. (10) is the ratio of the two, which reflects the comparison of the computing resource usage of the two strategies. By calculating the ratios of the two strategies in terms of time-to-solution and computational cost, they are $t_{LCODE}/t_{FW} \approx 35.59$ and $c_{LCODE}/c_{FW} \approx 1.0$, respectively, which shows that the method of using NN instead of simulation software is more advantageous in terms of time-to-solution. After the NN is trained, different objective functions can be used for optimization according to different research problems. When the number of research problems is greater than 1, the computational cost of using the framework will exceed 1, indicating that when the NN is used at least twice to optimize different problems, the computational cost will show an advantage. This is because training the neural network itself requires computational costs caused by data set support. After training the NN, the optimization of different research problems will not increase the computational cost of training the NN. If there are more than one research problems, the advantages of combining the optimization algorithm and NN in time-to-solution and computational cost will increase as the

TABLE 4. Parameters related to the speedup calculation.

Quantity	Symbol	Value
Time for calculating one event by LCODE	t_0	2.26 min
Time to train one neural network model	t_t	2.77 h
Time to predict once with the framework	t_p	0.44 s
Number of event for training NN model	n	2000
Number of hyperparameters to try	n_h	120
Number of CPU cores per evaluation by LCODE	r_0	72
Number of CPU cores to train and evaluate the framework	r_t	8
Number of generations	n_g	3000
Number of individuals per generation	n_i	1

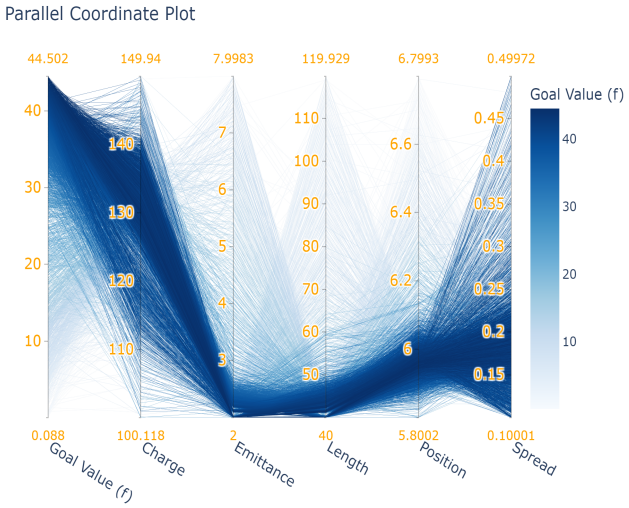

Fig. 6. During the BO process, the line graph of each generation's target f-value and the corresponding variables.

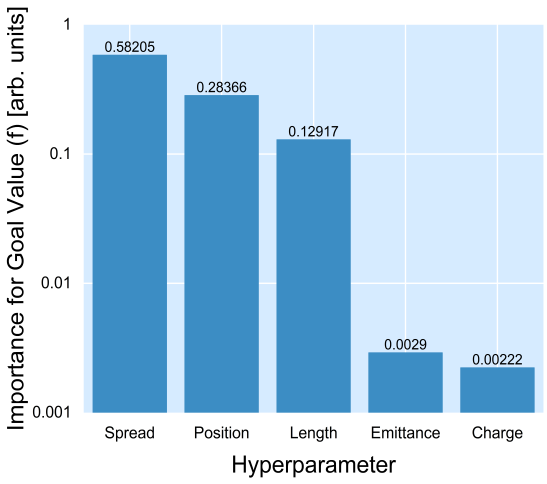

Fig. 7. Hyperparameter importance of BO.

number of research problems increases, as shown in Fig. 8.

$$t_{LCODE} = n_g \cdot t_0 \tag{5}$$

$$t_{FW} = t_0 + t_t + n_g \cdot t_p \tag{6}$$

$$\frac{t_{LCODE}}{t_{FW}} = \frac{n_g \cdot t_0}{t_0 + t_t + n_g \cdot t_p} \tag{7}$$

$$c_{LCODE} = n_g \cdot n_i \cdot t_0 \cdot r_0 \tag{8}$$

$$c_{FW} = n \cdot t_0 \cdot r_0 + n_h \cdot t_t \cdot r_t + n_g \cdot n_i \cdot t_p \cdot r_t \tag{9}$$

$$\frac{c_{LCODE}}{c_{FW}} = \frac{n_g \cdot n_i \cdot t_0 \cdot r_0}{n \cdot t_0 \cdot r_0 + n_h \cdot t_t \cdot r_t + n_g \cdot n_i \cdot t_p \cdot r_t} \tag{10}$$

VI. CONCLUSION

Our results have shown that for the AWAKE Run 2 optimization process, the combination of BO and NN can significantly reduce the computing time and the requirement for computing resource usage for finding the optimal solution under a certain computing accuracy. The data calculated by the simulation software LCODE is used as the training set of the NN, and a NN with 6 inputs and 4 outputs has been successfully trained. In the BO process, the prediction ability of the

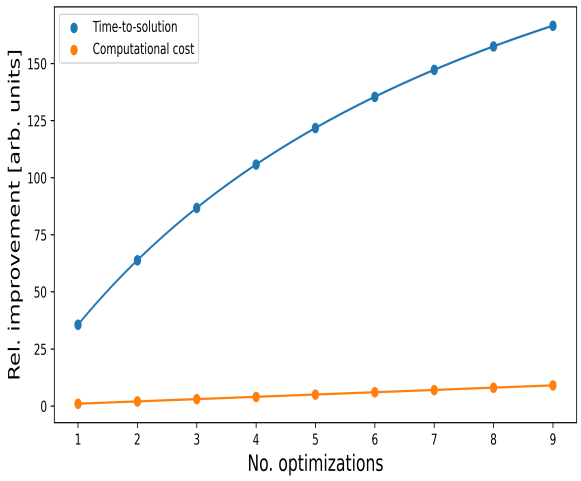

Fig. 8. Relative improvement in terms of time-to-solution and computational cost.

NN is used to replace the calculation process of the simulation software to find the optimal solution of the goal function. The results show that the framework reduces the optimization time (by speedup factor of 35.59), that is, about 35.59 times less computational time, which will increase if multiple optimizations need to be performed, because the training model is the main time-consuming and computationally intensive part. For more complex physical problems with more variables, it will be more advantageous to use NN instead of simulation software in the stage of BO. In this paper, first use of BO combined with NN to optimize PWFA, which demonstrates the great potential in the optimization of other particle accelerations, especially when the setup involves multiple variables.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request.

ACKNOWLEDGMENTS

This work was supported by the Major Science and Technology Project of Hubei Province(2021AFB001).

The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

[1] Tajima, T., Dawson, J. M., Laser Electron Accelerator. Phys. Rev. Lett. **43**, 267 (1979). doi: 10.1103/PhysRevLett.43.267

[2] Chen, Pisin, Dawson, J. M., Huff, Robert W. et al., Acceleration of Electrons by the Interaction of a Bunched Electron Beam with a Plasma. Phys. Rev. Lett. **54**, 693 (1985). doi: 10.1103/PhysRevLett.54.693

[3] Xinlu Xu, Fei Li, Frank S. Tsung et al., Generation of ultrahigh-brightness pre-bunched beams from a plasma cathode for X-ray free-electron lasers. Nature Communications **13**, 3364 (2022). doi: 10.1038/s41467-022-30806-6

[4] R. Pompili et al., Free-electron lasing with compact beam-driven plasma wakefield accelerator. Nature **605**, 659 (2022). doi: 10.1038/s41586-022-04589-1

[5] Labat, M., D. Alesini, M. P. Anania et al., Seeded free-electron laser driven by a compact laser plasma accelerator. Nat. Photon. **17**, 150 (2023). doi: 10.1038/s41566-022-01104-w

[6] Galletti, M., D. Alesini, M. P. Anania et al., Stable Operation of a Free-Electron Laser Driven by a Plasma Accelerator. Phys. Rev. Lett. **129**, 234801 (2022). doi: 10.1103/PhysRevLett.129.234801

[7] Wang, W., Ke Feng, Lintong Ke et al., Free-electron lasing at 27 nanometres based on a laser wakefield accelerator. Nature **595**, 516 (2021). doi: 10.1038/s41586-021-03678-x

[8] Wing M, Particle physics experiments based on the AWAKE acceleration scheme. Phil. Trans. R. Soc. A. **377**, 20180185 (2019). doi: 10.1098/rsta.2018.0185

[9] G. Xia, O. Mete, A. Aimidula et al., Collider design issues based on proton-driven plasma wakefield acceleration. Nucl. Instrum. Meth. A **740**, 173 (2014). doi: 10.1016/j.nima.2013.11.006

[10] W. Leemans, E. Esarey, Laser-driven plasma-wave electron accelerators. Physics Today **62**, 44 (2009). doi: 10.1063/1.3099645

[11] E. Adli, Towards a PWFA linear collider — opportunities and challenges. JINST **17**, T05006 (2022). doi: 10.1088/1748-0221/17/05/T05006

[12] Jerzy J. Jasielec, Bartłomiej Wierzba, Bartosz Grysakowski et al, Novel Strategy for Finding the Optimal Parameters of Ion Selective Electrodes. ECS Trans. **33**, 19 (2011). doi: 10.1149/1.3557871

[13] Michalewicz Z, Genetic algorithms + data structures = evolution programs. Springer-Verlag, New York (1992). doi: 10.1007/978-3-662-03315-9

[14] Katoch, S., Chauhan, S.S., Kumar, V, A review on genetic algorithm: past, present, and future. Multimedia Tools and Applications **80**, 8091 (2021). doi: 10.1007/s11042-020-10139-6

[15] Kennedy J, Eberhart RC, Particle swarm optimization. Proceedings of IEEE international conference on neural networks **4**, 1942 (1995). doi: 10.1109/ICNN.1995.488968

[16] Mockus, J, The Bayesian approach to global optimization. Control Inform. Sci **38**, 473 (1982). doi: 10.1007/BFb0006170

[17] Shahriari, B., Swersky, K., Wang, Z. et al., Taking the human out of the loop: a review of bayesian optimization. Proceedings of the IEEE **104**, 148 (2016). doi: 10.1109/JPROC.2015.2494218

[18] Jianren Shang, Production scheduling optimization method based on hybrid particle swarm optimization algorithm. J. Intell. Fuzzy Syst. **34**, 955 (2018). doi: 10.3233/JIFS-169389

[19] F. Irshad, S. Karsch, A. Döpp, Multi-objective and multi-fidelity Bayesian optimization of laser-plasma acceleration. Phys. Rev. Research. **5**, 013063 (2023). doi: 10.48550/arXiv.2210.03484

[20] E J Dolier, M King, R Wilson et al., Multi-parameter Bayesian optimisation of laser-driven ion acceleration in particle-in-cell simulations. New J. Phys. **24**, 073025 (2022). doi: 10.1088/1367-2630/ac7db4

[21] R.J. Shalloo, S.J.D. Dann, J.-N. Gruse et al., Automation and control of laser wakefield accelerators using Bayesian optimization. Nat Commun **11**, 6355 (2020). doi: 10.1038/s41467-020-20245-6

[22] Sören Jalas, Manuel Kirchen, Philipp Messner et al., Bayesian Optimization of a Laser-Plasma Accelerator. Phys. Rev. Lett. **126**, 104801 (2021). doi: 10.1103/PhysRevLett.126.104801

[23] Bellotti, R., Boiger, R., Adelmann, A, A. Fast, Efficient and Flexible Particle Accelerator Optimisation Using Densely Connected and Invertible Neural Networks. Information **12**, 351 (2021). doi: 10.3390/info12090351

[24] P Muggli, E Adli, R Apsimon et al., AWAKE readiness for the study of the seeded self-modulation of a 400 GeV proton bunch. Plasma Phys. Control. Fusion **60**, 014046 (2017). doi: 10.1088/1361-6587/aa941c

[25] R Assmann, R Bingham, T Bohl et al., Proton-driven plasma wakefield acceleration: a path to the future of high-energy particle physics. Plasma Phys. Control. Fusion **56**, 084013 (2014). doi: 10.1088/0741-3335/56/8/084013

[26] A. Caldwell, E. Adli, L. Amorim et al., Path to AWAKE: Evolution of the concept. Nucl. Instrum. Meth. A **829**, 3 (2016).

doi: 10.1016/j.nima.2015.12.050
[27] E. Gschwendtner, E. Adli, L. Amorim et al., AWAKE, the advanced proton driven plasma wakefield acceleration experiment at CERN. Nucl. Instrum. Meth. A **829**, 76 (2016). doi: 10.1016/j.nima.2016.02.026
[28] E. Gschwendtner, Konstantin Lotov, Patric Muggli et al., The AWAKE Run 2 Programme and Beyond. Symmetry **14**, 1680 (2022). doi: 10.3390/sym14081680
[29] Kumar, N., Pukhov, A., Lotov, K, Self-modulation instability of a long proton bunch in plasmas. Phys. Rev. Lett. **104**, 255003 (2010). doi: 10.1103/PhysRevLett.104.255003
[30] Schroeder, C.B., Benedetti,C., Esarey, E. et al., Growth and phase velocity of self-modulated beam-driven plasma waves. Phys. Rev. Lett. **107**, 145002 (2011). doi: 10.1103/PhysRevLett.107.145002
[31] A. Pukhov, N. Kumar, T. Tückmantel et al., Phase velocity and particle injection in a self-modulated proton-driven plasma wakefield accelerator. Phys. Rev. Lett. **107**, 145003 (2011). doi: 10.1103/PhysRevLett.107.145003
[32] E. Gschwendtner, AWAKE RUN 2 AT CERN. IPAC2021, Campinas, SP, Brazil, 1757-1760 (2021). doi: 10.18429/JACoW-IPAC2021-TUPAB159
[33] V. K. B. Olsen, E. Adli, P. Muggli, Emittance preservation of an electron beam in a loaded quasilinear plasma wakefield. Phys. Rev. Accel. Beams **21**, 011301 (2018). doi: 10.1103/PhysRevAccelBeams.21.011301
[34] R. Ramjiawan, S. Döbert, J. Farmer et al., Design and operation of transfer lines for plasma wakefield accelerators using numerical optimizers. Phys. Rev. Accel. Beams **25**, 101602 (2022). doi: 10.1103/PhysRevAccelBeams.25.101602
[35] A.P. Sosedkin, K.V. Lotov, LCODE: A parallel quasistatic code for computationally heavy problems of plasma wakefield acceleration. Nucl. Instrum. Meth. A **829**, 350 (2016). doi: 10.1016/j.nima.2015.12.032
[36] K.V. Lotov, Fine wakefield structure in the blowout regime of plasma wakefield accelerators. Phys. Rev. ST Accel. Beams **6**, 061301 (2013). doi: 10.1103/PhysRevSTAB.6.061301
[37] E. Adli, A. Ahuja, O. Apsimon et al., Acceleration of electrons in the plasma wakefield of a proton bunch. Nature **561**, 363 (2018). doi: 10.1038/s41586-018-0485-4
[38] M. Turner1, E. Adli, A. Ahuja et al., Experimental Observation of Plasma Wakefield Growth Driven by the Seeded Self-Modulation of a Proton Bunch. Phys. Rev. Lett. **122**, 054801 (2019). doi: 10.1103/PhysRevLett.122.054801
[39] E. Adli, A. Ahuja, O. Apsimon et al., Experimental Observation of Proton Bunch Modulation in a Plasma at Varying Plasma Densities. Phys. Rev. Lett. **122**, 054802 (2019). doi: 10.1103/PhysRevLett.122.054802
[40] John P. Farmer, Linbo Liang, Rebecca Ramjiawan et al., Injection tolerances and self-matching in a quasilinear wakefield accelerator. arXiv e-prints **2203**, 11622 (2022). doi: 10.48550/arXiv.2203.11622
[41] Linbo Liang, Guoxing Xia, Hossein Saberi et al., Simulation study of betatron radiation in AWAKE Run 2 experiment. physics.acc-ph arXiv **2204**, 13199 (2022). doi: 10.48550/arXiv.2204.13199